

Guiding Active Contours for Tree Leaf Segmentation and Identification^{*}

Guillaume Cerutti^{1,2}, Laure Tougne^{1,2}, Julien Mille^{1,3}, Antoine Vacavant⁴,
Didier Coquin⁵

¹ Université de Lyon, CNRS (firstname.lastname@liris.cnrs.fr)

² Université Lyon 2, LIRIS, UMR5205, F-69676, France

³ Université Lyon 1, LIRIS, UMR5205, F-69622, France

⁴ Clermont Université, Université d'Auvergne, ISIT, F-63001, Clermont-Ferrand
(antoine.vacavant@iut.u-clermont1.fr)

⁵ LISTIC, Domaine Universitaire, F-74944, Annecy le Vieux
(didier.coquin@univ-savoie.fr)

Abstract. In the process of tree identification from pictures of leaves in a natural background, retrieving an accurate contour is a challenging and crucial issue. In this paper we introduce a method designed to deal with the obstacles raised by such complex images, for simple and lobed tree leaves. A first segmentation step based on a light polygonal leaf model is first performed, and later used to guide the evolution of an active contour. Combining global shape descriptors given by the polygonal model with local curvature-based features, the leaves are then classified over nearly 50 tree species.

1 Introduction

In our everyday more urbanized and artificial world, the knowledge of plants, that used to constitute our most immediate environment, has somehow been lost, except for a handful of specialists. What is allegedly seen as unquestionable progress also scattered away the names and uses of so many trees, flowers and herbs. But nowadays, with a certain resurgence of the idea that plant resources and diversity ought to be treasured, the will to regain some touch with nature feels more and more tangible. And making it possible, for whoever feels the need, to identify a plant species, to learn its history and properties, is as much a way to transmit a vanished knowledge, as to allow people to get a glance at nature's unfathomable richness.

The identification of species is the first and essential key to understand the plant environment. Botanists traditionally rely on the aspect and composition of fruits, flowers and leaves to identify species. But in the context of a widespread non-specialist-oriented application, the predominant use of leaves, which are possible to find almost all year long, simple to photograph, and easier to analyze from two-dimensional images, is the most sensible and widely used approach in image processing. Considering the shape of a leaf is then the obvious choice to

^{*} This work has been supported by the French National Agency for Research with the reference ANR-10-CORD-005 (REVES project).

try to recognize the species. Our system intends then to classify a photograph of a leaf, which should be roughly centered and vertically-oriented as shown in Figure 1, over around 50 different tree species.

In this paper we present a method to classify simple and lobed leaves, using a polygonal modeling of leaf shapes and geometric botany-inspired descriptors. In Section 2 we will present related publications. Section 3 expounds the first segmentation step using a parametric active polygon, and Section 4 the refinement leading to the actual segmentation. The process of classification is then detailed in Section 5 and Section 6 relates the results of experiments both on white-background leaf images and natural scenes.

2 Related work

Plant recognition has recently been a subject of interest for various works. Few of them though consider acquiring images of leaves or flowers in a complex, natural environment, thus eluding most of the hard task of segmentation.

Nilsback and Zisserman [1] addressed the problem of segmenting flowers in natural scenes, by using a geometric model, and classifying them over a large number of classes. Saitoh and Kaneko [2] focus also on flowers, but in images with hard constraints on out-of-focus background. Such approaches are convenient for flowers, but lose much of their efficiency with leaves.

Many works on plant leaf recognition tend to avoid the problem by using a plain sheet of paper to make the segmentation easy as pie. Their recognition systems are then based on either statistical or geometric features: Centroid-Contour Distance (CCD) curve [3], moments [4,3], histogram of gradients, or SIFT points [1]. Some more advanced statistical descriptors, such as the Inner Distance Shape Context [5], or the Curvature Scale Space representation [6] that allows taking self-intersections into account, have also been applied to the context of leaf classification, while developed in a general purpose.

As a matter of fact, isolating green leaves in an overall not less green environment seems like a much tougher issue, and only some authors have designed algorithms to overcome the difficulties posed by a natural background. Teng, Kuo and Chen [7] used 3D points reconstruction from several different images to perform a 2D/3D joint segmentation using 3D distances and color similarity. Wang [4] performed an automatic marker-based watershed segmentation, after a first thresholding-erosion process. All these approaches are complex methods that seem hardly reachable for a mobile application. In the case of weed leaves, highly constrained deformable templates have been used [8] to segment one single species, providing good results even with occlusions and overlaps.

The concept of active contours, or snakes, have been introduced by Kass, Witkin and Terzopoulos [9] as a way to solve problems of edge detection. They are splines that adjust to the contours in the image by minimizing an energy functional. This energy is classically composed of two terms, an internal energy term considering the regularity and smoothness of the desired contour, and an external or image energy accounting for its adequation with the actual features in the image, based on the intensity gradient.

To detect objects that are not well defined by gradient, Chan and Vese [10] for instance based the evolution of their active contour on the color consistency of the regions, by using a level set formulation. Another region-based approach, relying this time on texture information, was proposed by Unal, Yezzi and Krim [11] who also introduced a polygonal representation of the contour.

But to include some knowledge about complex objects, a deformable template [12] can be used, with the asset of lightening the representation and storage space of the contour by the use of parameters. Felzenszwalb [13] represents shapes by deformable triangulated polygons to detect precisely described objects, including maple leaves, and Cremers [14] includes shape priors into level set active contours to segment a known object, but both approaches lack the flexibility needed to include knowledge about the shape of any possible leaf.

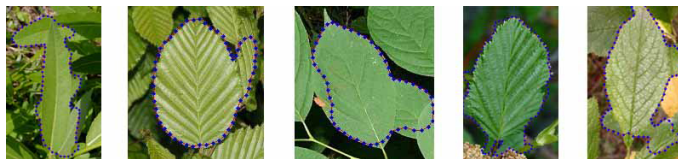


Fig. 1. Segmentation issues with unconstrained region-based active contours

In our case, the similarity between the background and the object of interest, and the difficulty to avoid adjacent and overlapping leaves constitute a prohibitive obstacle to the use of unconstrained active contours (Figure 1). The idea of using a template to represent the leaves is complicated by the fact that there is much more variety in shapes than for eyes or mouths. The only solution to overcome the aforementioned problems is however to take advantage of the prior knowledge we may have on leaf shapes to design a very flexible time-efficient model to represent leaves.

3 Parametric Polygonal Leaf Model

Even when considering trees only, leaves show an impressively wide variety in shapes. It is however necessary to come up with a representation of what a leaf is, that is accurate enough to be fitted to basically any kind of leaf. In a first time, we consider only simple leaves, including lobed and palmate ones, which covers already 90% of French broad-leaved tree species [15] and roughly the same proportion in all European species.

3.1 A polygon to describe leaf shapes

The general shape of a leaf is a key component of the process of identifying a leaf. Botanists have a whole set of terms describing either the shape of a simple leaf, of the lobes of a palmate leaf, or of the leaflets of a compound leaf. Examples of such shapes are displayed in Figure 3. The problem being that the borders between the different terms are not well defined, since leaves can naturally have non-canonical, intermediate shapes.

To sketch the shapes used in botany, we propose a light polygonal model based on a set of 4 parameters [16] The idea of defining a simple parametric

model to represent all these various shapes has the double advantage of turning an imprecise, blurry and quite subjective classification into numerical parameter values, and producing in the same time some descriptors accounting for the general shape of the leaf.

The chosen model relies on two points, base B and tip T , that define the main symmetry axis of the leaf. From this axis, we construct the 10 points defining the polygon, as displayed in Figure 2, using 4 numeric integer parameters:

- α_B , the opening angle at the base
- α_T , the opening angle at the tip
- w , the relative maximal width
- p , the relative position where this width is reached

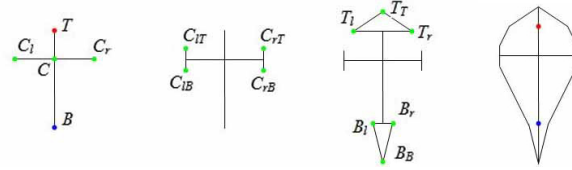


Fig. 2. Building the simple leaf model in 4 steps

This model is then able to cover correctly, with a very concise set of parameters, almost every shape used in botany. Figure 3 illustrates how, by changing manually the values of the parameters, it is possible to create shapes that visibly correspond to the different canonical shapes.

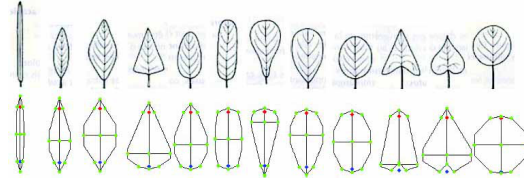


Fig. 3. Main leaf shapes and their corresponding hand-tuned models

The case of palmately lobed leaves is slightly more complex, but our hypothesis is that a palmate leaf can be seen as a superposition of several identical simple leaf shapes, joining at the base of the leaf. This is justified by the descriptions that can be found in flora books, where palmate leaves are characterized by the shape of their lobes using the same previously mentioned terms.

Concretely, we model a palmate leaf by overlaying an odd number of simple polygonal models, symmetrically arranged in pairs, and varying only in length and angle: a main lobe whose axis is defined by the points B and T , and several symmetric pairs of secondary lobes, whose axes all radiate from the actual base of the main lobe, the point B_B . Such representation implies the addition of a number of new parameters to make the construction represented in Figure 4 possible. In addition to the 4 parameters determining the shape of the lobes, we have to introduce:

- n_L , the number of pairs of lobes (1 means simple, 2 means 3-lobed)

- for each pair of lobes, $L(l)$, the relative length of the two symmetric lobes
- for each pair of lobes, $\alpha(l)$, the angle of the lobes with the main axis

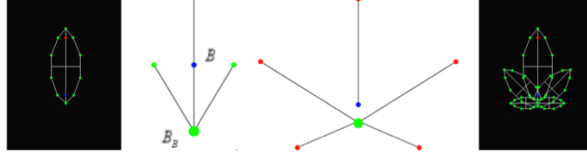


Fig. 4. Construction of the palmately lobed leaf model

The idea is then to apply a deformable template approach on this model, using the variations of the different parameters as elementary deformations, to make it fit best the leaf in the image. This has the major advantage of encapsulating some prior knowledge on the shape of the object, in the very definition of the model. The final polygon we obtain can later be used as a shape prior for a refined segmentation.

3.2 Color representation

A polygonal model inevitably being an approximation of the actual contour of the leaf, it is not relevant to base its evolution on the edge information, contained in the gradient. This is why we use only the color information to make the model fit the leaf in the image.

It is however impossible to come up with an *a priori* model for what the color of a leaf is, that would be accurate whatever the leaf, season and lighting. It is then absolutely necessary to estimate, for each new leaf we want to segment, the particular color model, which can only be performed with some rough idea of where the leaf lies in the image, implying constraints on its position. In the following, we assume that the leaf is roughly centered and vertically-oriented, so that an initialization of our template in the middle of the image contains almost only leaf pixels. Typically, a leaf in the image taken on purpose by a user will be large enough to ensure the correctness of the initialization.

We considered then that the color of a leaf can be modelled by a 2-component GMM estimated in the initial region, accounting for shaded and lighted or shiny areas, and defined by the parameters $(\mu_1, \sigma_1^2, \alpha_1)$ and $(\mu_2, \sigma_2^2, \alpha_2)$, respectively the mean, variance and weight of each gaussian distribution. Then the distance of a pixel x to the color model, in a 3-dimensional colorspace, is defined by a normalized 1-norm distance, written as following:

$$d(x, \mu_{1,2}, \sigma_1, \mu_2, \sigma_2) = \min_{g=1,2} \sum_{i=1}^3 \frac{|x_i - \mu_{g,i}|}{\sigma_{g,i}} \quad (1)$$

Based on this formulation, we can compute for every pixel in the image its distance to the color model, resulting in a map that measures the dissimilarity of pixels to the leaf, and where leaf pixels should appear in black and background pixels in gray-white. Based on the aspect of this map for different leaves and different color spaces, we chose to work in the $L^*a^*b^*$ colorspace, for which leaves were standing out best in distance maps.

3.3 Parametric Active Polygon

With an iterative process similar to active contours, the initial model will undergo a series of deformations, whose goal is to minimize an energy functional based on the afore described leaf dissimilarity map. The internal energy term that traditionally appears in the energy formulation can be considered as implicit here, as it is included in the construction rules of our model. The remaining external term is then expressed as:

$$E(\Omega) = \sum_{x \in \Omega} (d(x, \mu_1, \sigma_1, \mu_2, \sigma_2) - d_{max}) \quad (2)$$

The value d_{max} actually represents a balloon force, that will push the model to grow as much as possible. It can also be seen as a threshold, the distance to the color model for which a pixel will be costly to add into the polygon. The expected outcome is to produce the biggest region with as little leaf-dissimilar pixels as possible.

At each step, all the possible elementary variations of the parameters and of the base and tip points are examined, and the one leading to the greatest decrease of the energy is applied, until no deformation can bring it any lower. As such, this method presents the risk of getting stuck in local energy minima, this is why a heuristic close to simulated annealing is used. It is also necessary to constrain the model, that is basically flexible enough to take shapes that are not likely to ever be reached by leaves. To achieve that, actual leaf shapes were learned, and at every moment, the model has to remain within a certain distance to one of these reference shapes.

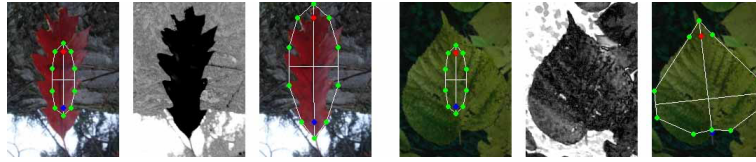


Fig. 5. Initialization, distance map and resulting polygon on two natural images

3.4 Collapsing lobes

The only parameter that has to be treated separately is the number of lobes n_L . Considering it on the same level as the other parameters does not make much sense, given the drastic changes in shape a modification of its value would induce. A way to solve the problem could have been to estimate in a preliminary step the number of lobes, and have this parameter fixed for the evolution of the model. But such an estimation, that would ultimately require some equivalent of a segmentation to be accurate, would just seem like going round in circles.

The method we kept treats the number of lobes in the same optimization process that fits the model to the leaf, but in a different way. After the color model has been estimated on the simple polygonal initialization shown in Figure 5 an excessive number of lobes is added to the model, typically resulting in a 11-lobed leaf model. The deformable template algorithm runs then freely on all the

parameters, except for the number of lobes. The expected result is that lobes will group together inside the actual lobes of the leaf. The key is then to eliminate the overlapping lobes until there remains only as many lobes as necessary.

This elimination step is performed between two temperature rises in the simulated annealing process. For each secondary pair of lobes, we evaluate an overlapping ratio with the previous one, expressed as:

$$r(l) = \frac{\frac{w}{2} \cdot L(l-1)}{\sin(\theta)} \cdot \frac{1}{L(l)} \cdot \cos(\theta) \quad (3)$$

where $\theta = \alpha(l) - \alpha(l-1)$. Figure 6 illustrates the origin of this formula where polygonal models are approximated by rectangles. This value can be seen as the ratio between the blue area (estimating the overlapped part of the considered lobe) and the whole lobe, weighed by a cosinus to favour more large angles between lobes. The actual result is the ratio between the lengths of the two red lines.

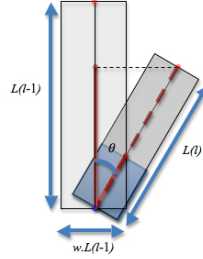


Fig. 6. Estimating the overlapping ratio between two lobes

Lobes that have a overlapping ratio greater than a given threshold (typically, 70%) are eliminated. In addition, a second pass is performed to ensure that the remaining lobes have a plausible repartition, which could correspond to a leaf. If a lobe makes a naturally unlikely angle with its predecessor (typically, more than 60 degrees) it is suppressed too. The result is that the perceived number of lobes will emerge from the very process that approximates the leaf by a model (as in Figure 7) and this even for the case of simple leaves, where all the lobes should eventually collapse into one.

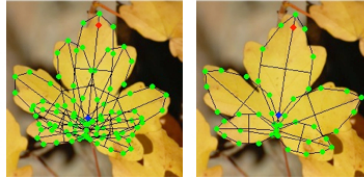


Fig. 7. Sample result of lobe suppression, before suppression and after evolution

The resulting model gives a good clue of what the shape of the leaf is, but it is not enough to carry out species identification, and the first resulting contour has to be later enhanced to hug the boundary of the leaf.

4 Guided Active Contour Segmentation

As initially mentioned, unconstrained active contours applied to the complex natural images we aim at dealing with would produce unsatisfying contours, that would try and make their way through every possible gap and flaw in the border of the leaf. The solution we propose is to use the polygonal model obtained after the first step not only as an initial leaf contour but also as a shape prior that will guide its evolution towards the real leaf boundary.

4.1 Energy formulation

We chose to use the active contour model defined by Kass et al. [9], including a guiding constraint and using leaf dissimilarity, both under the form of additional terms in the energy functional the contour strives to minimize. The general energy term, for a contour Γ delineating a region $\Omega(\Gamma)$, can be expressed as:

$$E(\Gamma) = \alpha E_{Leaf}(\Gamma) + \beta E_{Shape}(\Gamma) + \gamma E_{Gradient}(\Gamma) + \delta E_{Smooth}(\Gamma) - \omega E_{Balloon}(\Gamma) \quad (4)$$

Instead of having an external energy term based on color consistency, or distance to a mean, we decided to reuse the dissimilarity map from the previous step, considering we have already an efficient measure of how well a pixel should fit in the leaf, in terms of color. The corresponding energy term is then based on the dissimilarity function d detailed in Section 3 and can be written as:

$$E_{Leaf}(\Gamma) = \int_{\Omega(\Gamma)} d(x, \mu_1, \sigma_1, \mu_2, \sigma_2) dx$$

The guiding constraint term relies on a so-called stencil function K_{Shape} increasing with the distance to the polygonal contour Π , which is explicited in 4.2. It guarantees that contour points that are distant from the polygon that will be pushed back towards it, when neither the color nor the gradient would be strong enough to retain it. It has then the following form:

$$E_{Shape}(\Gamma) = \int_{\Gamma} K_{Shape}(\Gamma(s), \Pi) ds$$

However this is obviously not enough, and the contribution of the gradient computed on the image I is here essential, as it may allow the contour to stick to the actual boundaries of the leaf (which is ultimately the main objective) even when the color information would not be relevant enough. The gradient energy is expressed in order to penalize curve points located on pixels with low gradient magnitude and is simply formulated as:

$$E_{Gradient}(\Gamma) = \int_{\Gamma} -\|\nabla I(\Gamma(s))\| ds$$

Although the final contour has to be precise, to capture points and teeth on the leaf margin essentially, some moderate smoothing is still necessary to prevent the contour from being too noisy. And finally the balloon energy is here to counterbalance and stabilize the other energies by adding a constant force towards the outside of the contour:

$$E_{Smooth}(\Gamma) = \int_{\Gamma} \left\| \frac{d\Gamma}{ds} \right\|^2 ds \quad E_{Balloon}(\Gamma) = \int_{\Omega(\Gamma)} dx$$

The final variation of E is determined using calculus of variations, and the resulting evolution equation is implemented on a parametric curve following [17].

4.2 Guided by a stencil

The function we use is asymmetric with respect to the polygonal contour Π , as shown in Figure 8. The reason is that, our basic energy already using a balloon force, we want the force pushing the contour to the polygon to be stronger on the outside of the polygon, where it has to compete with the balloon, than on the inside, where the two resulting forces will push towards the same direction. We also introduce a margin Δ around the polygon where the contour should be free to evolve, *i.e.* where the shape energy is equal to 0. For a given pixel, located at a distance d from the polygonal contour, the cost function is then:

$$k(d) = \begin{cases} \max(d - \Delta, 0) & \text{on the inside of the polygon} \\ \max((d - \Delta)^2, 0) & \text{on the outside of the polygon} \end{cases} \quad (5)$$

A small improvement to this basic formulation is a slight elongation of the polygonal contour at the base and the tip of the leaf before the evaluation of the distance, resulting in the two round black areas on the map of Figure 8. This is justified by the fact that the shape of the leaf in these areas is really discriminant, and the contour absolutely has to be able to reach such determinant parts of the leaf. Since the polygonal model is by definition not well suited to account for local particular shapes, the contour needs a larger freedom around those crucial points.

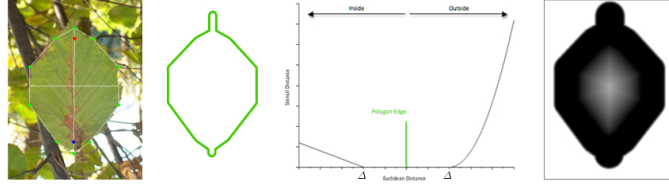


Fig. 8. Polygonal model, modified polygon, cost function, and stencil map

If we call Π^* the extension of the original polygonal contour Π at the base and at the tip, the final expression for our stencil function is: $K_{Shape}(x, \Pi) = k(d_{\Pi^*}(x))$ where $d_{\Pi^*}(x) = \min_{y \in \Pi^*} \|x - y\|$ represents the Euclidean distance of a pixel to the contour Π^* .

4.3 Ebb and flow

As such however, the various energies, though having all their role to play, are hard to harmonize into a coherent whole, efficient regardless of the image. Some of them may have conflicting interactions, the leaf dissimilarity term impeding for instance that the contour reaches the edge where the gradient term would make it stay. This is the reason why, rather than trying to find the optimal way of making those disparate forces cooperate smoothly, we decided to split the process into two phases where each will be able to express freely.

Starting from a contracted version of the polygon, to try to make sure that most of the points are located inside the leaf, the contour will undergo a first expansion step, where the coefficients of the balloon and gradient are preeminent, and the margin Δ is rather tolerant. The contour should swell with no concern for the color, sometimes getting outside of the leaf, but constrained to be blocked when it crosses strong gradients or when it gets quite far from the reference polygon.



Fig. 9. Initial polygon, contour after expansion, and final contour after contraction

The second phase has for purpose to bring back the contour from non-leaf areas where it might have wandered, with at the same time hanging on to the strong edges and not going back inside the leaf even when the dissimilarity would claim so. The coefficient of the leaf dissimilarity energy is this time very high while the balloon is reduced and enough gradient preserved to stick to the boundary of the leaf. The distance map to the polygon is recomputed with a smaller Δ , again in this perspective of shrinking the contour to the leaf and no further. Those two steps and the final contour we obtain are illustrated in Figure 9.

5 Geometric Features and Classification

Our purpose is then to classify new leaf images into one of the 50 species of non-compound-leaved trees of the ImageCLEF Plant Images Classification task database[18]. We made the choice of basing our classification on explicit morphological descriptors inspired by those used by botanists. These geometric criteria are those used to build the determination keys that have proved some skills when it comes to helping with species identification.

5.1 Global shape features

As stated earlier, the global shape of a leaf is one of the most relevant characteristics for its identification, even if it has to be completed by local details. We already presented a way of dealing with global shape features, by means of the parametric polygonal described in 3.1. As a matter of fact, the parameters resulting from the final evolution of polygon can directly be used as descriptors accounting for the global shape of the leaf, since it is actually the very purpose they were designed for.

Among them, we chose to keep only those who appeared to be the most discriminating ones: the four simple shape parameters w , p , α_B and α_T , and the

two lobe parameters for the first pair of secondary lobes $L(1)$ and $\alpha(1)$, these two being respectively set to 100 and 0 when the number of lobes is 1.

5.2 Contour characterization

The margin of the leaf is also a very important feature to spot. Its shape can be determining when trying to discriminate two species that have more or less the same global shape. It may consist of teeth of various sizes and frequencies, regularly arranged or not, from large spiny points, to small regular saw-like teeth, or even to a smooth entire border.

In order to cover these numerous possibilities, we based our contour descriptors by a measure of the curvature, using the osculating circle method [19,20] estimated at 1, 3, 5, and 10 points, for each point of the contour. Then we simply computed the means and variances of the resulting vectors, and used those 8 values to characterize how large and how regular the teeth of the margin are.

There is however a lot more to do to measure this crucial feature accurately, and if the descriptors we use are visibly performing satisfyingly, they are still very correlated, and do not account efficiently for the regularity of the teeth, or for doubly serrated leaves. Such basic measures also fail in representing the local shape at the base and the tip of the leaf, which are both essential elements for species identification.

5.3 Normalized classification

The learning base was built by extracting the aforementioned descriptors on 2101 white-background images of non-compound leaves of the ImageCLEF training database. Once computed, the base was normalized by adjusting all the values for each parameter, in order that their mean becomes 0 and their variance 1. Then for the k -th species among the 50 we treat, the specific mean and variance of each parameter are computed resulting in a model $\Phi_k = (\mu_k, \sigma_k)$ representing the species.

For a new image, the parameters extracted during and after the segmentation form a feature vector ϕ that has to be compared with every of the 50 species models. For each one of them we compute a normalized Euclidean distance in the parameters space, with the only particularity that the features are weighted differently, by a constant coefficient denoted by w_i for the i -th feature. The formulation of the distance is then: $d(\phi, \Phi_k) = \sqrt{\sum_i w_i \frac{(\phi(i) - \mu_k(i))^2}{\sigma_k(i)^2}}$

The species corresponding to the model Φ^* achieving the smallest distance is the one that is picked if we have to give one single answer. But when multiple answers are possible, we associate them with a confidence score, measuring how probable the species are. The only species that are to be displayed in a final list are the ones with a confidence score not equal to zero. If we designate by Φ^{**} the model corresponding to the second best distance, our recipe of a confidence measure is given by: $C(\phi, \Phi_k) = \max\left(1 - \frac{d(\phi, \Phi_k)}{d(\phi, \Phi^*) + d(\phi, \Phi^{**})}, 0\right)$

Our final classification result presents a list of species, in decreasing order with respect to the confidence we allow to each species.

6 Results

6.1 Segmentation results

To measure the accuracy of our segmentation process, we compare the binary mask representing the final contour with a hand-made binary segmentation of the same image. Segmentations were hand-performed on more than 200 photographs of the ImageCLEF Database[18]. Results were measured by computing an overlapping factor accounting for how well the region Ω our algorithm provides overlaps with the expected truth T : $\rho = \frac{|\Omega \cap T|}{|\Omega \cup T|}$

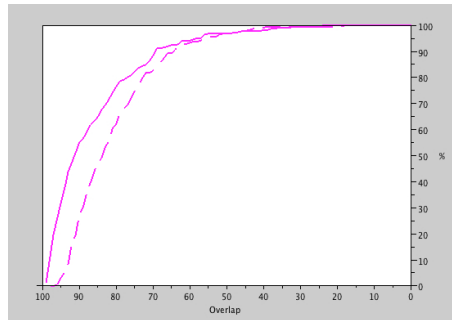


Fig. 10. Overlap curves for the polygonal model (dotted) and the final contour (plain)

Figure 10 displays the results for 226 segmentations, showing for each overlapping factor value the percentage of images in the database for which this score is reached. Both the polygonal model (dotted line) and the active contour result (plain line) are presented. The results show a mean overlap of 80,3% for the polygonal model and 85,8% for the final contour, with an overlapping score of 80% being reached for more than 78% of the images in the database.

6.2 Scan and pseudoscan images

As said previously, the training of our classifier was performed on these easier plain-background images. We used cross-validation during the learning base, by using 30% of the base as a testing base to get a reliable performance measure. Our typical classification scores for scan and pseudoscan images over 50 classes lie around 70% on the training base and 67% on the testing base, with respectively 97% and 94% of presence of the real class in the 5 first answers.

On the plant identification task proposed for ImageCLEF 2011 [18] our method achieved a classification score of $53.8 \pm 0.8\%$ on scan images and $52.6 \pm 1.8\%$ on scan-like images.

6.3 Natural background images

When using the same classifier for the photograph images of the training base, that contained only 26 of the 50 species, we reach only a score of 29%, with the

good answer being in the top 5 species in 63% of cases. This can of course be explained by the fact that returning one of 50 possible classes when there are actually only 26 increases the risk of making bad choices.

There is also the fact that we use the plain-background images as our reference, considering that the contours we obtained on these easier cases would be more trustworthy. However, by making this choice, we prevent our system from learning the defaults it may have on more complicated images, and consequently adapting to them, making it extremely less robust to the noisy, sometimes inaccurate contours obtained on such difficult pictures.

On the ImageCLEF plant identification task, we performed a score of 18.7 ± 6.4 % on such natural scene photographs.

7 Conclusions and Perspectives

We have presented a method designed to perform the segmentation of a leaf in a natural scene, based on the optimization of a polygonal leaf model used as a shape prior for an exact active contour segmentation. It also provides a set of global geometric descriptors that, later combined with local curvature-based features extracted on the final contour, make the classification into tree species possible.

The segmentation process is based on a color model that is robust to uncontrolled lighting conditions. But a global color model for a whole image may sometimes not be enough, for leaves that are not well defined by color only. The use of an additional texture model, or of an adaptive color model could lead to a good improvement.

There is also a need for better local contour descriptors, always keeping in mind what years of botany have found to be the most discriminative places to look at. These descriptors would also have to be more robust to the change from white-background images to real natural scenes, or to little flaws in the contour natural objects inevitably carry.

Nevertheless, it constitutes already a promising way of dealing with images of leaves in a complex background, giving promising results in terms of segmentation correctness, and proposing interesting descriptors relying on the criteria used by botanists to discriminate leaves.

Acknowledgements

We would like to thank **Hung Duong Viet** for the work on leaf segmentation he performed during his Master internship.

References

1. Nilsback, M.E., Zisserman, A.: Delving into the whorl of flower segmentation. In: British Machine Vision Conference. Volume 1. (2007) 570–579
2. Saitoh, T., Kaneko, T.: Automatic recognition of blooming flowers. International Conference on Pattern Recognition **1** (2004) 27–30

3. Wang, Z., Chi, Z., Feng, D., Wang, Q.: Leaf image retrieval with shape features. In: *Advances in Visual Information Systems*. Volume 1929 of *Lecture Notes in Computer Science*. (2000) 41–52
4. Wang, X.F., Huang, D.S., Du, J.X., Huan, X., Heutte, L.: Classification of plant leaf images with complicated background. *Applied Mathematics and Computation* **205** (2008) 916–926
5. Belhumeur, P., Chen, D., Feiner, S., Jacobs, D., Kress, W., Ling, H., Lopez, I., Ramamoorthi, R., Sheorey, S., White, S., Zhang, L.: Searching the world’s herbaria: A system for visual identification of plant species. In: *European Conference on Computer Vision*. (2008)
6. Mokhtarian, F., Abbasi, S.: Matching shapes with self-intersections: Application to leaf classification. *IEEE Transactions on Image Processing* **13** (2004) 653–661
7. Teng, C.H., Kuo, Y.T., Chen, Y.S.: Leaf segmentation, its 3d position estimation and leaf classification from a few images with very close viewpoints. In: *Proceedings of the 6th International Conference on Image Analysis and Recognition. ICIAR ’09* (2009) 937–946
8. Manh, A.G., Rabatel, G., Assemat, L., Aldon, M.J.: Weed leaf image segmentation by deformable templates. *Journal of agricultural engineering research* **80** (2001) 139–146
9. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. *International Journal of Computer Vision* **1** (1988) 321–331
10. Chan, T., Vese, L.: Active contours without edges. *IEEE Transactions on Image Processing* **10** (2001) 266–277
11. Unal, G., Yezzi, A., Krim, H.: Information-theoretic active polygons for unsupervised texture segmentation. *International Journal of Computer Vision* **62** (2005) 199–220
12. Yuille, A., Hallinan, P., Cohen, D.: Feature extraction from faces using deformable templates. *International Journal of Computer Vision* **8** (1992) 99–111
13. Felzenszwalb, P.: Representation and detection of deformable shapes. *PAMI* **27** (2004) 208–220
14. Cremers, D., Tischhäuser, F., Weickert, J., Schnörr, C.: Diffusion snakes: introducing statistical shape knowledge into the mumford-shah functional. *International Journal Of Computer Vision* **50** (2002) 295–313
15. Coste, H.: *Flore descriptive et illustrée de la France de la Corse et des contrées limitrophes*. (1906)
16. Cerutti, G., Tougne, L., Vacavant, A., Coquin, D.: A parametric active polygon for leaf segmentation and shape estimation. In: *7th International Symposium on Visual Computing*. (2011)
17. Mille, J.: Narrow band region-based active contours and surfaces for 2d and 3d segmentation. *Computer Vision and Image Understanding* **113** (2009) 946–965
18. Goau, H., Bonnet, P., Joly, A., Boujemaa, N., Barthelemy, D., Molino, J.F., Birnbaum, P., Mouysset, E., Picard, M.: The clef 2011 plant images classification task. (2011)
19. Vacavant, A., Coeurjolly, D., Tougne, L.: A framework for dynamic implicit curve approximation by an irregular discrete approach. *Graphical Models* **71** (2009) 113–124
20. Hermann, S., Klette, R.: A comparative study on 2d curvature estimators. In: *International Conference on Computing: Theory and Applications*. (2007) 584–589